

What is claimed is:

- 1 1. A method comprising:
2 receiving an informational database request; and
3 determining whether a result set corresponding to said informational database
4 request is stored in a cache, and if so, returning said result set in response to said
5 informational database request, and if not,
6 sending said informational database request to a database, wherein said
7 database generates said result set, and
8 determining whether to add said result set to said cache with reference to
9 the cache-worthiness of said result set.
- 1 2. The method of claim 1, wherein said cache stores one or more result sets and
2 the database requests corresponding to said one or more result sets, and wherein
3 said determining whether said result set is stored in said cache comprises
4 comparing said database request with said database requests stored in said
5 cache.
- 1 3. The method of claim 2, wherein said result set is determined to be stored in said
2 cache if said received database request is identical to one of said database
3 requests stored in said cache.
- 1 4. The method of claim 2, wherein said database request comprises a parameter
2 and a query string, and wherein said result set is determined to be stored in said
3 cache if said parameter and said query string are identical to the parameter and
4 query string corresponding to one of said database requests stored in said cache.
- 1 5. The method of claim 2, wherein said result set is determined to be stored in said
2 cache if said database request is logically the same as one of said database
3 requests stored in said cache.
- 1 6. The method of claim 1, further comprising:
2 receiving a transactional database request, wherein said transactional database
3 request targets one or more objects within said database; and

invalidating one or more result sets stored in said cache that include data from said one or more objects.

7. The method of claim 1, further comprising:
collecting cache-worthiness data for said result set; and
determining a cache-worthiness value with reference to said cache-worthiness data, wherein said cache-worthiness value is reflective of the cache-worthiness of said result set.

8. The method of claim 7, wherein said cache-worthiness data includes data that is reflective of a number of times said result set has been returned in response to informational database requests.

9. The method of claim 8, wherein said cache-worthiness data further includes data that is reflective of the size of said result set.

10. The method of claim 9, wherein said cache-worthiness data further includes data that is reflective of the amount of time required for said database to generate said result set.

11. The method of claim 10, wherein said cache-worthiness data further includes data that is reflective of the number of times said result set has been invalidated.

12. The method of claim 10, further comprising degrading said cache-worthiness data.

13. The method of claim 12, wherein said degrading comprises degrading said cache-worthiness data on a timed basis.

14. The method of claim 12, wherein said degrading comprises degrading said cache-worthiness data responsive to a miss on said cache.

15. The method of claim 7, wherein said cache-worthiness data includes time data that is reflective of the average time to execute and fetch said result set, and wherein said method further comprises degrading said time data according to

$$\text{avgTime} = \text{avgTime} + ((\text{newTime} - \text{avgTime}) / \text{hit})$$

wherein avgTime is an average time to execute and fetch said result set, newTime is the most recent measurement of the time to execute and fetch said result set, and hit is the current count of the number of times said result set is requested.

16. The method of claim 7, wherein said cache-worthiness value is determined according to:

$$\text{cache-worthiness value} = (\text{hit} / \text{invalid} + 1) * \text{time}$$

wherein hit is the number of times said result set is requested, invalid is the number of times the result set is invalidated, and time is the average time required to execute and fetch said result set.

17. A result set cache comprising:

first program code means to receive an informational database request; and

second program code means to determine whether a result set corresponding to said informational database request is stored in the result set cache, and if so, to return said result set in response to said informational database request, and if not,

to send said informational database request to a database, wherein said database generates said result set, and

to determine whether to add said result set to the result set cache with reference to the cache-worthiness of said result set.

18. The result set cache of claim 17, further comprising a memory to stores one or more result sets and the database requests corresponding to said one or more result sets, and wherein said second program code means comprises program code means to compare said database request with said database requests stored in said cache.

19. The result set cache of claim 18, wherein said result set is determined to be stored in said cache if said database request is identical to one of said database requests stored in said cache.

1 20. The result set cache of claim 18, wherein said database request comprises a
2 parameter and a query string, and wherein said result set is determined to be
3 stored in said cache if said parameter and said query string are identical to the
4 parameter and query string corresponding to one of said database requests
5 stored in said cache.

1 21. The result set cache of claim 18, wherein said result set is determined to be
2 stored in said cache if said database request is logically the same as one of said
3 database requests stored in said cache.

1 22. The result set cache of claim 17, further comprising:
2 third program code means to receive a transactional database request, wherein
3 said transactional database request targets one or more objects within said database; and
4 fourth program code means to invalidate one or more result sets stored in said
5 cache that include data from said one or more objects.

1 23. The result set cache of claim 17, further comprising:
2 fifth program code means to collect cache-worthiness data for said result set;
3 and
4 sixth program code means to determine a cache-worthiness value with reference
5 to said cache-worthiness data, wherein said cache-worthiness value is reflective of the
6 cache-worthiness of said result set.

1 24. The result set cache of claim 23, wherein said cache-worthiness data includes
2 data that is reflective of a number of times said result set has been returned in response
3 to informational database requests.

1 25. The result set cache of claim 24, wherein said cache-worthiness data further
2 includes data that is reflective of the size of said result set.

1 26. The result set cache of claim 25, wherein said cache-worthiness data further
2 includes data that is reflective of the amount of time required for said database to
3 generate said result set.

1 27. The result set cache of claim 26, wherein said cache-worthiness data further
2 includes data that is reflective of the number of times said result set has been
3 invalidated.

1 28. The result set cache of claim 17, further comprising seventh program code
2 means to degrade said cache-worthiness data.

1 29. The result set cache of claim 28, wherein said seventh program code means
2 executes on a timed basis.

1 30. The result set cache of claim 28, wherein said seventh program code means
2 executes responsive to a miss on the result set cache cache.

1 31. A result set cache comprising:
2 memory to store one or more result sets and metadata associated with each of
3 said result sets;

4 first program code means to collect cache-worthiness data associated with said
5 one or more result sets;

6 second program code means to determine a cache-worthiness value for each of
7 said one or more result sets, wherein said cache-worthiness values are determined with
8 reference to said cache-worthiness data; and

9 third program code means to update the contents of the result set cache based at
10 least in part on said cache-worthiness values.

1 32. The result set cache of claim 31, wherein said third program code means
2 updates the contents of the result set cache on a timed basis.

1 33. The result set cache of claim 31, wherein said third program code means
2 updates the contents of the result set cache on an as-needed basis.

1 34. The result set cache of claim 31, wherein said one or more result sets comprise
2 result sets that have been completely fetched.

1 35. The result set cache of claim 31, wherein said third program code means uses a
2 cache victimization strategy to update the contents of the result set cache.

1 36. The result set cache of claim 31, further comprising fourth program code means
2 to degrade said cache-worthiness data.

1 37. The result set cache of claim 36, wherein said third program code means
2 updates the contents of the result cache responsive to said fourth program code means
3 degrading said cache-worthiness data.

1 38. The result set cache of claim 36, wherein said third program code means and
2 said fourth program code means are executed responsive to a database request being
3 received by the result set cache.

1 39. The result set cache of claim 36, wherein said third program code means and
2 said fourth program code means are executed responsive to a miss on the result set
3 cache.

1 40. The result set cache of claim 36, wherein said third program code means and
2 said fourth program code means are executed on a timed basis.

1 41. The result set cache of claim 31, further comprising fifth program code means
2 to remove result sets stored in said memory responsive to said third program code
3 means.

1 42. The result set cache of claim 41, wherein said fifth program code means uses a
2 best-fit algorithm to remove result sets stored in said memory.

1 43. A system comprising:
2 a database;
3 an application;
4 a result set cache including:
5 first program code means to receive an informational database request
6 from said application; and

second program code means to determine whether a result set corresponding to said informational database request is stored in said result set cache, and if so, to return said result set to said application in response to said informational database request, and if not,

to send said informational database request to said database, wherein said database generates said result set, and

to determine whether to add said result set to said result set cache based at least in part on the cache-worthiness of said result set.

44. The system of claim 43, further comprising a cache driver, wherein said application calls said cache driver to send said informational database request to said result set cache.

45. The system of claim 44, further comprising a database driver, wherein said result set cache calls said database driver to send said informational database request to said database.

46. The system of claim 43, wherein said application and said result set cache use client-side resources, and wherein said database uses server-side resources.

47. The system of claim 43, wherein said application uses client-side resources, and wherein said result set cache and said database use server-side resources.

48. The system of claim 43, wherein said application uses client-side resources, said database uses server-side resources, and said result set cache comprises a stand-alone appliance.

49. A system comprising:
a plurality of clients;
a database;
a result set cache configured to store result sets, wherein said result sets are generated by said database in response to database requests issued by said clients, and wherein said result sets are stored separately for each client.

1 50. The system of claim 49, wherein a key is associated with each of said result
2 sets, and wherein said key is generated with reference to a client identifier and a
3 query string.

1 51. The system of claim 49, wherein said client identifier comprises actual user
2 information.

1 52. The system of claim 49, wherein said result set cache restricts access to said
2 result sets such that said clients are only able to access their own separately
3 stored result sets.

1 53. The system of claim 49, wherein said result set cache comprises first program
2 code means for translating database requests received from said clients into
3 canonical form, wherein a key is associated with each of said result sets, said
4 key being generated in canonical form with reference to a client identifier and a
5 query string.

1 54. A method of maintaining consistency in a result set cache, wherein said cache
2 stores one or more result sets, said method comprising:
3 receiving a database request;
4 determining whether said database request is informational or transactional; and
5 if said database request is transactional, invalidating result sets stored in the
6 result set cache that include data targeted by said database request.

1 55. The method of claim 54, wherein each of said result sets is generated based on
2 one or more objects stored in a database, and wherein said invalidating
3 comprises:
4 parsing said database request to determine whether any of said one or more
5 objects are affected by said database request; and
6 marking those results sets that were generated based on the affected objects.

1 56. The method of claim 54, further comprising invalidating result sets stored in the
2 result set cache on a timed basis.